

Beyond black and white: the impact of Asian peers on scholastic achievement

Article (Published Version)

d'Este, Rocco and Einiö, Elias (2021) Beyond black and white: the impact of Asian peers on scholastic achievement. *Economics of Education Review*, 83. a102129 1-13. ISSN 0272-7757

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/98809/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.



Beyond Black and White: The Impact of Asian Peers on Scholastic Achievement[☆]

Rocco d'Este^{a,*}, Elias Einiö^b

^a Department of Economics, University of Sussex, United Kingdom

^b VATT Institute for Economic Research, Finland

ARTICLE INFO

JEL Classifications:

I21

I24

J15

Keywords:

Peer effects

Education

Public schools

Racial composition

Test scores

ABSTRACT

This paper examines the effects of Asian peers on non-Asian student achievement in New York City public schools. We use exogenous variation in the share of Asian students across cohorts within schools stemming from a fertility shock among the Asian population in the Chinese year of the Dragon. Results show that a 10-percentage-point increase in the share of Asian students reduces non-Asian math and ELA scores by 0.14 and 0.16 standard deviations. The reduction in achievement is associated with an increase in the share of non-Asian students who fail to demonstrate the skills expected at the grade, especially in math.

1. Introduction

Since the famous 1954 Supreme Court case *Brown vs. Board of Education*, which deemed segregation in public schools unconstitutional, extensive school desegregation efforts have been undertaken in the U.S. public school sector (Welch & Light, 1987). A central motivation for these policies has been the presumption that school racial composition directly affects student achievement (Hanushek, Kain, & Rivkin, 2009). Much of the previous research on this issue has focused on assessing the impacts of the share of black students on educational outcomes. However, the number of non-black minority students has grown in many U.S. school districts in recent decades, but very little is known about the consequences of this change (Rivkin & Welch, 2006). In this paper, we examine the impacts of Asian students on their peers' academic performance. Asian students are a particularly interesting minority group because they perform very well in school (e.g., Chao & Tseng, 2002). Their importance has also been increasing in recent years because Asian Americans are the fastest-growing major racial group in the U.S. (Budiman, Cilluffo, & Ruiz, 2019).

The effects of Asian students are not obvious *ex ante*. Positive impacts may occur if they help their schoolmates or aid learning through

questions and answers (e.g., Hanushek, Kain, Markman, & Rivkin, 2003). They may also be less disruptive and require less attention in class, allowing teachers to devote more time and resources toward other students (Lazear, 2001). However, other students could also be discouraged by the high level of achievement of Asian peers. This may lead them to exert less effort, lowering their scholastic performance (Rogers & Feller, 2016). Furthermore, teachers could respond to changes in student composition by adjusting the pace and coverage of instruction, which may be beneficial for some students but harm others (e.g., Duflo, Dupas, & Kremer, 2011).

The objective of this study is to provide causal estimates of the effects of Asian peers on their schoolmates' test scores. The key empirical challenge arises from the potential selection of Asian students into schools. Asian parents have often high expectations for their children's education (Chao & Tseng, 2002). Therefore, they may choose better schools, where peer quality could also be higher. As a result, the share of Asian students could be correlated with unobserved school and peer quality, inducing a positive bias in regressions of student achievement on the share of Asian students.

We address these issues by exploiting exogenous variation in the share of Asian students due to the common belief among the Asian

[☆] We are grateful to Jordi Blanes i Vidal, Maria De Paola, Michael Gilraine, Sandra McNally, Hessel Oosterbeek, Francis Kramarz, Victor Lavy, Devin Pope, Fabian Waldinger, Roberto Weber, Ulf Zölitz and several seminar participants for helpful comments and suggestions.

* Corresponding author.

E-mail addresses: r.d-este@sussex.ac.uk (R. d'Este), elias.einio@vatt.fi (E. Einiö).

population that children born in the Chinese year of the Dragon are luckier and brighter than those born under other zodiac signs. This belief generates considerable shocks to fertility in the Asian populations in the Dragon years.¹ Our empirical strategy is based on the fact that the relative magnitude of these fertility shocks varies geographically with the local size of the Asian population. In areas with a small historical share of the Asian population, the fertility shock in a Dragon year induces only small differences in the share of Asian children between cohorts, while in areas with a large historical share of the Asian population, it results in a disproportionately large proportion of Asian children in the Dragon cohort compared to other cohorts.

Our study covers 1080 public elementary and middle schools in New York City (NYC), which has one of the largest Asian populations among U.S. metropolitan areas (around 14% of the population). We use NYC Department of Education (DOE) data on average math and English language arts (ELA) test scores in third through eighth grade by school and race in the academic years 2005/2006 through 2011/2012. We geocode the schools by address and link them to 1990 census tract-level population data to measure the historical population structure in a school's neighborhood. In our analysis, the key Asian groups are those influenced by the Chinese culture. For this reason, we base our preferred instrument on the historical local share of the Chinese population, but we also show that results are broadly similar when we use instruments based on wider Asian groups.

We first show that an increase of eight percentage points in the historical Chinese population share leads to one additional Asian student in a school-grade-year cell in the Dragon cohort compared to other cohorts. Because the number of non-Asian students is little affected, the number of Asian students in the Dragon cohort increases disproportionately in areas with high historical share of the Chinese population. In line with this, our first-stage estimates show that a 10-percentage-point increase in the historical Chinese population share induces a 1.5-percentage-point increase in the *share* of Asian students in the Dragon cohort compared to other cohorts within a school, which corresponds to a 10% increase from the sample mean.

Our main finding is a statistically significant negative impact of the share of Asian students on non-Asian achievement. Our preferred estimates suggest that a 10-percentage-point increase in the share of Asian students reduces non-Asian math scores by around 0.14 standard deviations (henceforth σ) and non-Asian ELA scores by around 0.16 σ . Moreover, our analysis of the effects of Asian peers on the share of students in four performance levels reveals a reduction in non-Asian proficiency in both subjects. We find an increase in the share of non-Asian students at the lowest performance level in math, who fail to demonstrate the skills expected at the grade. For ELA, the results show a reduction in the share of non-Asian students at the highest performance level, while the impacts are weaker at the lower tail of the performance distribution. These differences in the impacts lead us to argue that the possible mechanisms generating the negative impacts on non-Asian achievement, such as student discouragement and teacher responses, are likely to be subject-specific.

We account for possible congestion effects by controlling for grade-level enrollment in our regression analysis. Our estimates also hold up to an extensive battery of robustness checks. In particular, we show that neighborhood-specific trends correlated with the non-Asian test scores are unlikely to be a source of bias in our analysis, and that changes in class size and teacher resources do not drive our estimates. Moreover, we show that excluding schools where the fertility shock may most likely trigger student mobility, such as those located nearby a charter or

private school or in areas with especially high levels of Chinese exposure, does not affect our results. These findings suggest that selection of non-Asian students in response to the fertility shock cannot explain the reduction in achievement. We also show that our results are very similar when adjacent cohorts, which could be affected by between-grade social spillovers, are excluded from the sample.

Our study contributes to the literature that examines the effects of school racial composition on educational outcomes. Previous studies have mainly focused on the impacts of the share of black students. Angrist and Lang (2004) examine the impacts of the Metropolitan Council for Educational Opportunity (Metco) desegregation program that moved mainly black students to better schools in Boston. They find little evidence of socially or statistically significant effects of Metco students on their non-Metco classmates. Hanushek et al. (2009) exploit patterns of racial composition for cohorts of students as they age within Texas public schools, finding that the share of black students adversely affects test scores with larger effects on black than on white students. Hoxby (2000) employs idiosyncratic variation across cohorts within schools in Texas. She finds that black students reduce test scores of their peers. Her study also estimates the effects of Asian students, but the precision of the estimates is low. This is likely due to the small share of Asian students in the Texas data (less than 3%). We extend this literature by providing novel evidence on the impacts of Asian student share on test scores in a setting that combines data from a school district with a large Asian population and significant quasi-experimental variation in the share of Asian students. To our knowledge, our paper is the first to provide quasi-experimental evidence on the impacts of Asian peers on their schoolmates' scholastic achievement.

The paper proceeds as follows. Section 2 provides details of the institutional background and presents data sources and descriptive statistics. Section 3 documents the shock on the share of Asian students in the Dragon cohort and discusses the estimation strategy. Section 4 provides the main results and robustness checks. Section 5 discusses the magnitude of the estimates and possible mechanisms. Section 6 concludes.

2. Data

Our study covers 1080 public elementary and middle schools in New York City. This section gives an overview of our data. Further details and links to publicly available files are provided in Appendix A.

The key components of our dataset are publicly available files provided by the DOE, which is one of the largest schooling authorities in the U.S., serving around 1.1 million students. The DOE files include school-level information on the group means of ELA and math test scores by race/ethnicity for third through eighth grade in the academic years 2005/2006 through 2011/2012.² The math tests cover the following topics: (1) number sense and operations, (2) algebra, (3) geometry, (4) measurement, and (5) statistics and probability. Tests in the earlier grades emphasize basic content, such as number sense and operations, whereas tests in the later grades focus on more advanced topics, such as algebra and geometry. The ELA tests are designed to assess students in three learning standards: (1) information and understanding, (2) literary response and expression, and (3) critical analysis and evaluation. The ELA tests include multiple-choice, short-response, reading, and listening exercises, as well as brief editing tasks. The number of correct answers in a test is converted by the DOE into a "scale score" which is the main

¹ Mocan and Yu (2017) show that births spike in the Dragon years 2000 and 2012 in China. Johnson and Nye (2011) provide evidence of the Dragon effect for the 1976 cohort among Asian immigrants in the United States. Yip et al. (2002) document the Dragon effect in Hong Kong for cohorts born in 1976, 1988, and 2000.

² The tests are administered in the spring semester. The DOE provides the ELA and math mean scores when the number of students in a school-grade-year-race/ethnicity cell is larger than five.

outcome of our analysis.³

The DOE files also include the share of students at four performance levels. Score thresholds for these levels are determined annually at the state level. At level 1, a student fails to demonstrate the skills expected at the grade. At level 2, learning standards are only partially met. Students in level 3 meet the learning standards expected at the grade, while students at level 4 demonstrate a thorough understanding of the topics and meet the learning standards with distinction. These data allow us to examine the impacts of Asian peers on non-Asian achievement distribution across the four performance levels.

The DOE group mean test score files contain information on the number of students taking the math and ELA tests that we use to calculate enrollment and racial/ethnic shares. Math and ELA tests are obligatory. Therefore, the number of students taking the tests can be expected to be close to actual enrollment. If attendance in the math and ELA tests within a school-grade-year-ethnicity/race cell is not equivalent, we use the larger value. We have assessed the accuracy of test attendance as a measure of enrollment by first aggregating it at the school-year level and then comparing it to the corresponding enrollment figures drawn from New York State School Report Cards. The correlation between these variables is 0.98, which indicates that test attendance is an accurate measure of enrollment. To reduce the effect of outliers in the regression analysis, we cap enrollment at the upper tail to the 97.5th percentile of the distribution. Lastly, the DOE provides information on average class size (at the school-grade-year level) from the academic year 2006/2007 onward.

Annual School Report Cards provide school addresses, which we use to geocode schools. Our primary source for school coordinates is the U.S. Census Batch Geocoder. We manually check the resulting address matches and coordinates, and manually geocode schools for which the Census Batch Geocoder does not provide coordinates. We are able to assign coordinates to around 99% of the schools.⁴ School Report Cards also provide information on the number of teachers and teachers with less than three years of experience at the school-year level. We use these data as control variables in our robustness analysis.

We construct variables for the historical population structure in a school's neighborhood using 1990 tract-level census data by ethnicity/race and boundary shapefiles provided by the Minnesota Population Center. We implement a geographic information system (GIS) procedure to find census tracts within 500 m of the school. When several nearby census tracts are identified, we use the population-weighted average of the population shares. We call the area covered by these nearby census tracts the "school neighborhood". The 1990 Chinese population share in a school neighborhood is our primary measure of historical Chinese exposure. We also report results using other measures based on wider definitions of Asian groups (e.g., all Asians and Asians excluding Asian Indians) and various geographic scopes (census tracts within 1000, 2000, and 3000 m).

Summary Statistics. We restrict the sample to school-grade-year-race/ethnicity cells for which both math and ELA mean test scores are observed. Thus, our baseline sample is the same for both math and ELA outcomes.⁵ Table 1 provides summary statistics. Panel A displays means

Table 1
Summary Statistics.

	Full Sample		N
	Mean	Standard Deviation	
Panel A. Test Scores			
Math, All	675.1	(23.0)	59,399
Math, Non-Asian	670.6	(20.7)	49,872
Math, Asian	701.2	(17.1)	9527
Math, White	692.4	(17.8)	9141
Math, Black	664.2	(18.1)	18,824
Math, Hispanic	667.5	(18.3)	21,907
ELA, All	658.6	(16.5)	59,399
ELA, Non-Asian	656.2	(15.6)	49,872
ELA, Asian	672.4	(14.5)	9527
ELA, White	674.1	(14.6)	9141
ELA, Black	653.0	(12.7)	18,824
ELA, Hispanic	652.1	(13.3)	21,907
Panel B. Enrollment, Student Shares, and Class Size			
Enrollment	110.66	(72.87)	23,996
Asian Student Enrollment	16.4	(30.2)	23,996
Black Student Enrollment	33.4	(34.9)	23,996
Hispanic Student Enrollment	45.09	(44.05)	23,996
White Student Enrollment	14.9	(21.29)	23,996
Asian Student Share	11.84	(17.8)	23,996
Black Student Share	34.3	(30.7)	23,996
Hispanic Student Share	39.9	(27.2)	23,996
White Student Share	13.9	(21.5)	23,996
Class Size	24.6	(4.5)	19,949

Notes: Panel A shows means and standard deviations for group-mean test score data at the school-grade-year-ethnicity/race level for 1080 New York City public schools. The data cover the years 2006–2012 and include grades 3–8. The number of observations is lower for subgroups because group means are available only for schools where more than five students are observed in a school-grade-year-ethnicity/race cell. The means in Panel A are weighted by the number of students taking the test. Panel B shows unweighted means and standard deviations for data on enrollment and share of students by race at the school-grade-year level. Enrollment is measured by the number of students taking the ELA and math tests (when attendance in the two tests is not equal, the larger value is used). Class size is the average within a school-grade-year cell and available from the year 2007 onwards.

and standard deviations of the group mean test scores for all students and by race, weighted by the number of students taking the test. The means for the math and ELA scores are 675.1 and 658.6. These are equal to the means in the underlying individual-level test score distributions. On the other hand, standard deviations for individual-level test score distributions cannot be recovered from the group means. Therefore, when assessing the magnitude of our estimates, we use the DOE statistics on standard deviations retrieved from individual-level data in New York State. The average standard deviation for grades three through eight is 40.5 for ELA and 41.9 for math.⁶

The test score means by race show that Asian students perform better than black and Hispanic students in both math and ELA. They also perform better than white students in math. In ELA, the mean score for Asian students is slightly lower than for white students. In both ELA and math, Asian students have the largest share at the highest performance level (around 48% in math and 12% in ELA, see Appendix Table A1). Asian students have also the smallest share at the lowest performance level (around 2.5% in math and 4.4% in ELA). Overall, Asian students perform well in both subjects and are particularly high achieving in math.

Appendix Figure A2 shows group-mean test score distributions by

³ The aim of the DOE scaling procedure is to improve the comparability of the test scores across grades and years. The DOE considers the difficulty of the question, its capacity to differentiate between high- and low-performing students, and the likelihood of getting a correct answer by guessing. The scaling by grade has no impact on our econometric analysis, because we include grade-by-year fixed effects in all regressions. Also, the tests scores are not standardized to have a common standard deviation. Therefore, the test scores of Asian students do not mechanically affect the test scores of non-Asian students.

⁴ This leaves us with a baseline sample of 1,080 schools. Appendix Figure A1 shows an example of geocoded schools and census tracts in Manhattan.

⁵ This restriction reduces the main estimation sample by 0.3% (142 observations). Results are very similar when this restriction is not imposed and are available upon request.

⁶ These statistics are drawn from "New York State Testing Program 2006: English Language Arts, Grades 3–8" and "New York State Testing Program 2006: Mathematics, Grades 3–8", Tables 41 and 46. We use the average of grade-level standard deviations weighted by the number of grade-level observations.

race/ethnicity and for Asian Dragon and non-Dragon students. Appendix Table A2 shows a formal test of differences in mean test scores between Asian Dragon and non-Dragon students. In a specification controlling for school fixed effects, the Asian students in the Dragon cohort have 0.813 points higher mean ELA score and 3.162 points higher mean math score. We reject the null hypothesis of equivalent group means, suggesting that Asian Dragon students perform better, on average, compared to Asian non-Dragon students. The shape of the distributions is similar, however, and the distributions have common support across a wide range.⁷

Panel B of Table 1 shows descriptive statistics for enrollment and shares of students by race. The mean enrollment at the grade level is around 110 students. The average class size is 24.6 students. The mean Asian student share is 11.8% with a standard deviation of 17.8%. The mean 1990 Chinese population share based on census tracts within the 500-meter radius is 2.7% with a standard deviation of 6.3% (Appendix Table A3).

3. Empirical strategy

3.1. Asian fertility shock in the dragon year

In the Chinese calendar, the Dragon year appears once every 12 years. According to a widespread belief among many East Asian cultures, children born in these years are luckier, brighter, and more likely to flourish. This belief generates fertility shocks in populations among which it is prevalent. Previous research finds spikes in birth rates in the Dragon years in China (Mocan & Yu, 2017), many East Asian regions (e.g., Goodkind, 1995; Yip, Lee, & Cheung, 2002) and among the Asian population in the U.S. (Johnson & Nye, 2011). According to the U.S. Census Bureau data, Asian births per 1000 individuals are around 7.5% higher in 2000, compared to the average rate in the years 1998–1999 and 2001–2002, while the non-Asian birth rate does not show a similar spike (see Appendix Figure A3).

Our empirical strategy employs a shock to the Asian birth rate in the Chinese Dragon year, which started on the 5th of February in 2000 and ended on the 23rd of January in 2001. The increase in the Asian fertility rate in this Dragon year mainly affects births in the Western calendar year 2000. Hence, we use the terms “2000 cohort” and “Dragon cohort” interchangeably. We account for the fact that the Chinese Dragon year continued in the first 23 days of the Western calendar year 2001 by excluding the 2001 cohort in a robustness specification. This does not affect our results appreciably.

Fig. 1 shows growth rates in total enrollment of Asian and non-Asian students from cohort $t - 1$ to cohort t in our data. There is a dramatic spike in Asian enrollment in the 2000 cohort; while there is little between-cohort deviation before the 2000 cohort, enrollment of Asian students increases by around 10% from the 1999 to 2000 cohort and declines by around 5% from the 2000 to 2001 cohort.⁸ We do not observe a similar spike in non-Asian enrollment. This implies that the fertility shock increases the *share* of Asian students in the Dragon cohort. Our empirical strategy exploits the fact that this shock is disproportionately larger in areas with large historical Asian population. Next, we provide a formal discussion of this relationship.

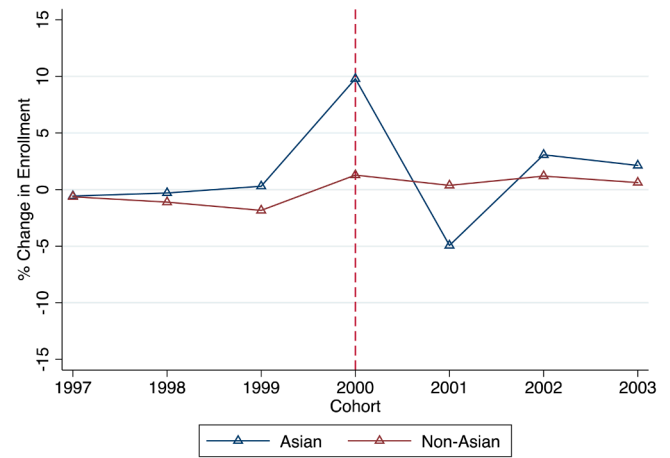


Fig. 1. Growth Rate of Enrollment of Asian and Non-Asian Students in New York City Public Schools by Birth Cohort.

Notes: The figure is based on annual enrollment of a cohort, averaged across calendar years. The growth rate is calculated as the percentage change in enrollment from the previous cohort. Data source: New York City Department of Education (2006–2012).

3.2. Geographic variation in the fertility shock

Consider a school neighborhood with A Asian and H non-Asian births in a cohort born in a non-Dragon year. The share of Asian children in a non-Dragon cohort is then $a = A/(A + H)$. Suppose that the Dragon year increases Asian births by $\delta \cdot 100\%$ and has no impact on the number of non-Asian births. The share of Asian children in the Dragon cohort will then be $a_D = (1 + \delta)A/((1 + \delta)A + H)$. It is straightforward to show that the increase in the share of Asian children between the Dragon and non-Dragon cohorts, $a_D - a$, is the following function of the size of the fertility shock in the Dragon year and the share of Asian children in the non-Dragon cohort:

$$g(a, \delta) = \frac{a - a^2}{\delta^{-1} + a}. \quad (1)$$

This function is concave and nonnegative when $a \in [0, 1]$ and has a maximum at $a = 0.5$ for $\delta > 0$. This relationship implies that a fertility shock of $\delta = 0.075$ induces a difference of around 1.8 percentage points in the share of Asian children in the Dragon cohort compared to other cohorts between areas with $a = 0$ and $a = 0.5$.

Fig. 2 shows the function $g(a, 0.075)$ for $a \in [0, 1]$ and the empirical distribution of the share of Asian students in the non-Dragon cohorts born one to three years before the Dragon cohort (in 1997–1999). For the majority of school neighborhoods, the share of Asian students in these non-Dragon cohorts is within the range where $g(a, 0.075)$ is increasing. As a result, the average derivative of $g(a, 0.075)$ evaluated across the distribution of the share of Asian students in these cohorts is positive (0.055). It is worth noting that the nonlinear relationship in Eq. (1) could affect our analysis. However, our results are very similar when we allow for nonlinearity in our empirical model.

3.3. IV estimation

In our analysis, the key Asian groups are those influenced by the Chinese culture. For this reason, we base our preferred instrumental variable on the local Chinese population share in 1990. NYC is an especially suitable metropolitan area for our research design because the 1990 Chinese population share varied considerably across neighborhoods, as shown in Fig. 3. The figure also shows that schools in our data are scattered across areas with high and low historical Chinese exposure.

We exploit variation in the share of Asian students induced by the

⁷ It is important to note that we control for cohort fixed effects in our regression analysis. Therefore, these between cohort differences in Asian achievement do not affect our estimates. We also show below that our instrument does not have a statistically or economically significant impact on average Asian achievement.

⁸ In levels, the increase from the 1999 to 2000 cohort is 910 Asian students. We note the negative growth rate for the 2001 cohort is due to the number of Asian students declining back toward the pre-Dragon levels. In absolute terms, this drop is smaller than the increase in 2000. Therefore, we cannot test whether the effect is different in the case of a negative exogenous shock to the share of Asian students.

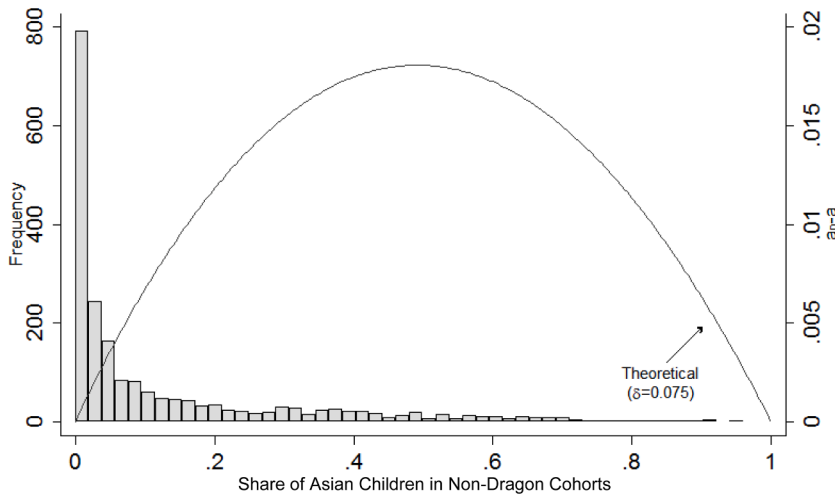


Fig. 2. Theoretical Impact of the Fertility Shock in the Dragon Year.

Notes: This figure displays the theoretical impact of a 7.5-percentage-point fertility shock in a Dragon year on the share of Asian children in the Dragon cohort compared to other cohorts (right axis). The histogram shows the empirical distribution of the share of Asian students in the third grade for three non-Dragon birth cohorts born in 1997–1999 (left axis).

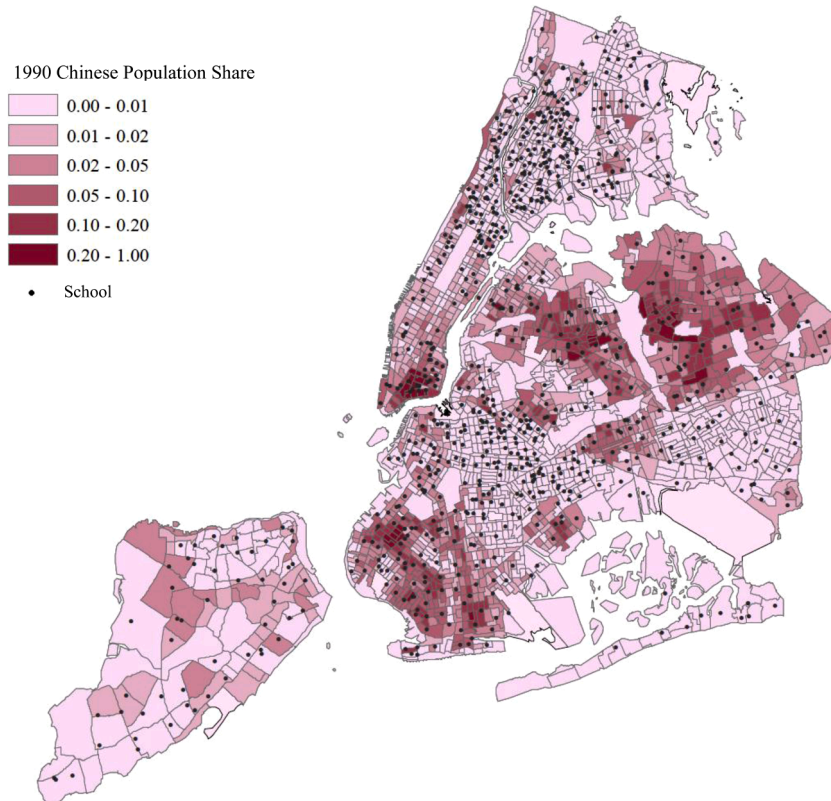


Fig. 3. Chinese Population Share by Census Tract in New York City, 1990.

Notes: This figure displays the Chinese population share by census tract in New York City in 1990. Dots represent public schools in our data. Data sources: U.S. Census Bureau 1990 census data and 1990 census tract shapefiles, both provided by the Minnesota Population Center; New York State School Report Cards.

disproportionately large fertility shock in the Dragon year in areas with a high historical Chinese population share by estimating the following two-stage least squares (TSLS) model:

$$y_{rsqt} = \rho_1 CS_{s,1990} + \rho_2 Dragon_{gt} + \gamma AS_{sqt} + \beta_1' X_{rsqt} + u_{rsqt} \quad (2a)$$

$$AS_{sqt} = \tau_1 CS_{s,1990} + \tau_2 Dragon_{gt} + \tau_3 CS_{s,1990} \times Dragon_{gt} + \beta_2' X_{rsqt} + \nu_{rsqt} \quad (2b)$$

where y_{rsqt} is the mean test score of non-Asian students in racial/ethnic group r in school s , grade g , and year t . AS_{sqt} is the share of Asian students

in school s , grade g , and year t . $Dragon_{gt}$ is a binary indicator taking the value one if the cohort born in 2000 is in grade g in year t and zero otherwise. $CS_{s,1990}$ is the Chinese population share in 1990 in the neighborhood of school s . X_{rsqt} is a vector of control variables. In our baseline specification, we include enrollment at the grade, and school, race, and grade-by-year fixed effects.⁹ We weight the regressions by the number of students taking the test and cluster the standard errors at the

⁹ Note that when school and grade-by-year fixed effects are included the terms for main effects of $CS_{s,1990}$ and $Dragon_{gt}$ become redundant.

census-tract level.

The IV model uses the interaction term between the Dragon cohort dummy and the 1990 Chinese population share in a school's neighborhood as an instrument for the share of Asian students. The first-stage coefficient on the instrument, τ_3 , recovers the difference in the within-school effect of the fertility shock on the share of Asian students in the Dragon cohort compared to other cohorts between schools in neighborhoods with high and low historical Chinese population shares.

As many previous studies examining the impacts of racial composition in schools, our strategy uses within school variation in the share of Asian students across cohorts. The key difference is that, rather than using a large number of idiosyncratic shocks as a source of identifying variation, we exploit an explicit population shock due to a cultural belief.¹⁰ This allows us to use variation in the share of Asian students stemming from the historical differences in the local population structure, which, conditional on school fixed effects, is plausibly exogenous with respect to scholastic achievement realized almost two decades later when the students in our data enter the school.

The key identifying assumption of this empirical strategy is that, conditional on enrollment at the grade and school, race, and grade-by-year fixed effects, the population shock due to the Dragon belief is uncorrelated with unobservable factors that affect both the share of Asian students and the non-Asian test scores. To lend credibility to this assumption, we show that our results are little affected when we control for year and year of birth trends interacted with the 1990 Chinese population share. This indicates that neighborhood-specific trends correlated with non-Asian test scores are unlikely to be a major source of bias in our analysis.

Another threat for the causal interpretation of our IV estimates is the potential movement of non-Asian children across schools due to the larger local Asian cohort, which may mechanically affect school-level non-Asian test score distributions. However, as argued by Carrell, Hoekstra, and Kuka (2018), changing schools is a rather extreme response to negative peer effects, because it likely involves moving residence.¹¹ Moreover, our robustness analysis shows that excluding schools that may be the most susceptible to student mobility, such as those located nearby a charter or private school or in areas with especially high levels of Chinese exposure, does not affect our results. We also find no evidence of within-school attrition of non-Asian students due to the additional Asian students in the Dragon cohort.¹² Because we control for grade-level enrollment in all regressions, our results are unaffected by the impact of the Asian fertility shock on the number of students in the grade, which can affect scholastic outcomes through congestion.¹³

¹⁰ Because we use an explicit population shock as a source of variation in student composition, our empirical strategy is also linked to Imberman, Kugler, and Sacerdote (2012), who employ quasi-experimental variation in peer composition arising from explicit shocks to the local population structure due to Hurricanes Katrina and Rita to estimate the impacts of evacuee students on educational outcomes of non-evacuee students.

¹¹ They examine the impacts of disruptive peers and find negative peer effects on test scores that are of a similar magnitude as ours.

¹² We also note that the causal interpretation of our estimates is not affected by avoidance of classes with peers who are negatively affecting test scores, because our regressions identify the average grade-level impact of the share of Asian students within schools. Such avoidance may occur if, for instance, parents lobby school principals to move their children to classes with fewer peers that have negative impacts on their child's achievement (see e.g., Carrell, Hoekstra, and Kuka, 2018).

¹³ One might be still concerned that congestion in pre-schools could affect our results. We do not have test score data at the pre-school level that would allow us to test for this hypothesis directly. However, because we do not find evidence of congestion in primary schools affecting our results, and the correlation between pre-school and primary school enrollment is high (0.93), pre-school congestion is unlikely to explain our results.

Our model is a group-mean version of the standard linear-in-means peer regression where the peer characteristic is a dummy for being Asian. Because this variable is pre-determined and not affected by peer interaction, our estimates are not affected by the reflection problem (Manski, 1993). Our analysis is also not biased by the potential mechanical correlation between the peer mean of the Asian dummy and the own value for the Asian dummy (see e.g., Angrist, 2014), because we focus on the outcomes of non-Asian students. Lastly, given that Asian background is likely measured with high accuracy, our estimates are unlikely to be affected by measurement error in the peer characteristic, which has been shown to attenuate estimates in quasi-experimental settings (Feld & Zölitz, 2017).

3.4. Reduced-Form effects on enrollment, class size, and asian achievement

We start the empirical analysis by examining the impact of the instrument on grade-level enrollment of Asian, black, white, and Hispanic students to demonstrate how it affects the number and composition of students. We also estimate the impacts on class size and Asian test scores to understand whether they are affected by the additional Asian students.

Table 2 reports the effect of the instrument on grade-level enrollment by subgroup, grade-level average class size, and Asian test scores in the third grade (Panel A), which is the earliest grade that we observe in our data, and in all grades three through eight (Panel B), which are included in our baseline sample. In Panel A, the instrument has a positive and statistically significant effect on enrollment of Asian students in the third grade ($p < 0.05$). The estimate implies that an increase of around 8.4 percentage points in the 1990 Chinese population share in a school's neighborhood induces one additional Asian student in the Dragon cohort compared to other cohorts. This effect corresponds to an increase of around 6.1% from the sample mean of Asian enrollment. The estimates for enrollment of black, white, and Hispanic students are all statistically insignificant. In Panel B, showing the results for grades 3–8, the impact is positive for Asian enrollment and insignificant for other ethnic groups, except for Hispanic students for whom the estimate is negative and significant at the 10% risk level. We note that a negative coefficient is expected as we condition on total enrollment in the grade; when total enrollment is fixed, a positive shock to the number of students in one group means that the number of students in other groups needs to reduce, on aggregate. Nevertheless, one might still worry that the larger coefficient for Hispanic students compared to white or black students could mean that our empirical strategy does not identify the impact of the change in the share of Asian students alone, but also of a change in the composition of the non-Asian group. To address this, below we also estimate IV specifications controlling for the composition of non-Asian students. Reassuringly, this turns out to have little impact on our results. Furthermore, Appendix Table A4 shows that the instrument has no significant impacts on the shares of white, black, Hispanic, and female students. Overall, the findings in Table 2 indicate that the instrument has a significant positive impact on Asian enrollment. The similar findings for the third grade and all grades indicate that the impact of the instrument on the number of Asian students does not vanish when the Dragon cohort moves across grades.

In column 5 of Table 2, the impact of the instrument on class size is small and insignificant for both grade 3 and all grades. This result stems likely from the fact that class size is limited to 25 students in NYC public schools. Therefore, an increase in total enrollment does not need to result in larger class size, because classes that exceed the threshold are likely to be split.

Columns 6 and 7 of Table 2 show the impact of the instrument on Asian test scores. This regression allows us to assess whether the additional Asian students in the Dragon cohort are similar in terms of scholastic performance, on average, compared to Asian students who would be born in the Dragon year in the absence of the Dragon belief.

Table 2
Effect of the Instrument on Enrollment, Class Size, and Asian Test Scores.

	(1) Enrollment Asian	(2) Black	(3) White	(4) Hispanic	(5) Class Size	(6) Asian Test Score Math	(7) ELA
Panel A. Grade 3							
1990 Chinese population share (%) * Dragon	0.119** (0.0507)	−0.0332 (0.0372)	0.0201 (0.0279)	−0.0698 (0.0444)	−0.000557 (0.0136)	0.0254 (0.0446)	0.0212 (0.0401)
Dependent Mean	14.57	28.32	15.09	39.33	22.31	704.86	676.16
Observations	4919	4919	4919	4919	4117	1971	1971
Number of schools	724	724	724	724	713	328	328
Panel B. All Grades							
1990 Chinese population share (%) * Dragon	0.156*** (0.0541)	−0.00397 (0.0253)	−0.0216 (0.0157)	−0.0957* (0.0529)	0.00685 (0.0129)	0.00991 (0.0239)	0.0113 (0.0274)
Dependent Mean	16.40	33.43	15.72	45.09	24.59	701.25	674.10
Observations	23,996	23,996	23,996	23,996	19,942	9478	9478
Number of schools	1080	1080	1080	1080	1071	542	542

Notes: This table reports estimates of the effect of the instrument on grade-level enrollment, grade-level average class size, and Asian test scores. We report results for grade 3, which is the earliest grade that we observe in our data, and for grades 3–8, which are included in our baseline sample. The instrument is the interaction between the 1990 Chinese population share and the Dragon dummy, equal to one for the Dragon cohort and zero otherwise. The number of observations is lower in column 5 than in columns 1–4 because average class size is missing for some observations in the baseline sample. Estimations in columns 6–7 use a sample in which the Asian test scores are observed. All specifications control for grade-level enrollment and school and grade-by-year fixed effects. Standard errors clustered at the census-tract level are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 3
Effects of Asian Peers on Non-Asian Test Scores.

	(1) First Stage	(2) Reduced Form	(3) IV	(4) OLS
Panel A. Non-Asian ELA Score				
1990 Chinese population share (%) × Dragon	0.154*** (0.0568)	−0.101** (0.0415)		
Asian student share (%)			−0.656** (0.288)	0.0199 (0.0415)
Panel B. Non-Asian Math Score				
1990 Chinese population share (%) × Dragon	0.155*** (0.0570)	−0.0906** (0.0419)		
Asian student share (%)			−0.583** (0.291)	0.0585 (0.0384)

Notes: This table reports coefficients from IV and OLS regressions of ELA and math scores of non-Asian students on the share of Asian students. Outcomes are mean test scores by school, grade, year, and race/ethnicity. The sample includes 49,872 observations (1080 schools). The instrument is the interaction between the 1990 Chinese population share and the Dragon dummy, equal to one for the Dragon cohort and zero otherwise. All specifications control for grade-level enrollment and school, grade-by-year, and race fixed effects. Columns 1 and 2 display the first-stage and reduced-form coefficients on the instrument. Columns 3 and 4 display the IV and OLS coefficients on the share of Asian students. All regressions are weighted by the number of students taking the test. Standard errors clustered at the census-tract level are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Estimates for both ELA and math are small and insignificant, indicating that the instrument does not change the average Asian peer achievement.

In Appendix Table A5, we examine the impact of the instrument on the share of Asian students in each of the four performance levels. We note that the performance level shares alone do not reveal the magnitude of differences in student achievement because both small and large changes in the test score can move a student from one level to another.¹⁴ Therefore, it is important to interpret the results in the context of the estimated impacts on Asian test scores in Table 2. The results for Asian performance level shares suggest that the instrument increases the share of Asian students at the lowest performance level 1. However, as shown above, this does not result in lower average achievement in either subject. One potential explanation is that the test score differences that generate the differences in the performance level shares are small.

Another explanation is that an increase in the share of well-achieving students offsets the negative impact on the average test scores of the increase in the share of students at the lowest performance level. Indeed, we do find positive, although statistically insignificant, impacts of the instrument on the shares of students at the highest performance level 4. For math, the point estimate for the level 4 share is more than three times larger compared to the point estimate for the level 1 share, but it is less precisely estimated.

Overall, the results in this section indicate that the fertility shock in the Dragon year increases the number of Asian students and has little impact on class size. While the fraction of students at the highest and lowest performance levels appears to be higher among the additional Asian Dragon students, their average test scores are not significantly different from other Asian students in the Dragon cohort.

¹⁴ A marginal change in the test score can move a student from one level to another if the student is close to a level cutoff.

4. Results

4.1. Effects on test scores

Table 3 shows our baseline IV estimates. The first-stage estimates are around 0.15 and significant at the 1% risk level.¹⁵ These estimates imply that, within a school, the share of Asian students is 1.5 percentage points higher in the Dragon cohort compared to other cohorts in areas with 10 percentage points higher 1990 Chinese population share. The reduced-form effects on non-Asian test scores are negative and statistically significant for both ELA and math. The IV estimates of the impact of the share of Asian students on non-Asian test scores are -0.656 for ELA and -0.583 for math, and both are significant at the 5% risk level.¹⁶ These estimates mean that a 10-percentage-point increase in the share of Asian students reduces non-Asian ELA and math scores by around 6.6 and 5.8 points, or by around 0.16σ and 0.14σ , respectively.¹⁷

We also follow the recommendation of Andrews, Stock, and Sun (2019) and calculate the Anderson-Rubin (AR) confidence intervals, which are robust against weak identification, using the Stata *weakiv* package. We note that, in the single-instrument setting, the TSLS estimator is approximately unbiased even under weak identification (see e. g., Angrist & Pischke, 2008; Skeels & Windmeijer, 2018). However, the size of the TSLS test based on the conventional t-statistic may be distorted (e.g., Stock and Yogo, 2005). The robust AR confidence interval is not affected by such distortions and should therefore be used to check for robustness of confidence intervals against weak instruments (Andrews et al., 2019). Reassuringly, the 95% robust AR confidence intervals for our IV coefficient on the share of Asian students are

$[-1.801, -0.230]$ for ELA and $[-1.671, -0.116]$ for math, suggesting that the rejection of the null hypothesis is not driven by weak identification. This is reassuring given that first-stage F-statistics are both below the rule of thumb of 10 (7.293 for ELA and 7.386 for math).

The fourth column in Table 3 shows the corresponding OLS estimates. These estimates are positive, small, and insignificant for both subjects.¹⁸ The upward bias in the OLS estimation is in line with the existence of unobserved factors that vary over time within schools and that are positively correlated with both the share of Asian students and non-Asian student achievement. For instance, Asian parents, who can have high expectations for their children's education (see e.g., Chao & Tseng, 2002; Mocan & Yu, 2017), may decide to enroll their children into schools where educational outcomes improve. This can induce a positive bias in the OLS regressions.

4.2. Internal validity

Table 4 shows results for several robustness specifications that examine the validity of our IV approach. Panel A replicates the baseline IV estimates for comparison.

In Panel B, we control for linear terms of calendar year and birth year interacted with the 1990 Chinese population share. This increases the

¹⁵ The first-stage coefficients are slightly different for math and ELA because the regressions are weighted by the number of students taking the test, which is not always the same for both subjects. Appendix Figure A4 shows the distribution of the residual share of Asian students from a regression controlling for grade-level enrollment and school, grade-by-year, and race fixed effects. The figure shows that there is substantial variation left in the share of Asian students after controlling for these fixed effects and enrollment. Appendix Figure A5 displays a scatter plot of predicted Asian share on observed Asian share. The relationship is fairly linear with a slope (standard error) of 0.019 (0.00075).

¹⁶ For a specification that does not control for enrollment, the IV estimates (standard errors) are -0.740 (0.294) for ELA and -0.618 (0.293) for math.

¹⁷ We use the DOE statistics on standard deviations in New York State. These are 40.5 for ELA and 41.9 for math. See section 2 for details.

¹⁸ The OLS estimates are also positive, and become statistically significant, in a specification controlling for school-specific time trends (Appendix Table A6).

Table 4
Specification Checks.

	(1) Non-Asian ELA Score	(2) Non-Asian Math Score	(3) Number Of Schools	(3) Observations
Panel A. Baseline	-0.656^{**} (0.288)	-0.583^{**} (0.291)	1080	49,872
Panel B. Year and Cohort Trends	-0.543^* (0.313)	-0.773^* (0.418)	1080	49,872
Panel C. Controls	-0.754^{**} (0.360)	-0.715^{**} (0.355)	1076	41,235
Panel D. Controls, replace levels with shares	-0.519^* (0.277)	-0.522^{**} (0.260)	1074	41,147
Panel E. Exclude Most Exposed Areas	-2.175^{***} (0.719)	-0.842 (0.516)	899	40,941
Panel F. Exclude Schools Nearby a Private or Charter School	-0.756^{**} (0.351)	-0.730^{**} (0.352)	888	41,958
Panel G. Exclude Adjacent Cohorts	-0.837^{***} (0.317)	-0.633^{**} (0.298)	1080	39,230
Panel H. Grades 3–6	-0.589^{**} (0.300)	-0.557^* (0.335)	1077	37,074
Panel I. Quadratic Instrument	-0.688^{**} (0.326)	-0.600^{**} (0.283)	1080	49,872

Notes: Panel A reports the baseline IV estimates. Panel B includes calendar year and birth year interacted with the 1990 Chinese population share as control variables. Panel C controls for average class size, black enrollment, white enrollment, and Hispanic enrollment at the grade level, and for pupils per teacher ratio and share of teachers with fewer than three years of experience at the school level. Panel D is otherwise similar as panel C but replaces enrollment variables with the shares of students in the corresponding subgroups and adds the share of female students. Panel E excludes schools in the top quintile of the 1990 Chinese population share. Panel F excludes schools in the bottom quintile of the distance to a private or charter school. Panel G excludes cohorts born one year before or after the Dragon cohort. Panel H shows results for grades 3–6, in which we observe the Dragon cohort. Panel I adds the square of the 1990 Chinese population share interacted with the Dragon dummy as an instrument. All specifications control for grade-level enrollment and school, grade-by-year, and race fixed effects. Regressions are weighted by the number of students taking the test. Standard errors clustered at the census-tract level are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

effect to -0.773 for math and reduces it to

-0.543 for ELA. While controlling for trends leads to a lower precision of the estimation, both estimates are significant at the 10% risk level.

Panel C controls for average class size, enrollment of black, white, and Hispanic students, number of students per teacher, and the fraction of teachers with fewer than three years of experience. The coefficient on the share of Asian students is -0.754 for ELA ($p < 0.05$) and

-0.715 for math ($p < 0.05$), suggesting that class size, changes in enrollment of the non-Asian subgroups, and teacher resources are not driving the results. The specification in Panel D is otherwise similar as Panel C but replaces the enrollment variables with the shares of students in the corresponding non-Asian subgroups and adds the share of female students in the grade. The estimate for math (-0.522 ; $p < 0.05$) is only slightly lower than in the baseline specification. The estimate for ELA reduces to -0.519 , but it is still significant at the 10% risk.

One potential confounding factor could be mobility of students across schools as a response to the Asian fertility shock. The Asian Dragon cohort could increase competition for available slots in high-

quality schools in high-exposure areas, pushing some poorly performing non-Asian students into other schools. This would mechanically increase average non-Asian test scores in high-exposure areas. As discussed in Section 3, the potential movement of students across schools is unlikely to be large enough to generate our estimates. Moreover, attrition of low-performing non-Asian students in high-exposure areas would go against us finding a negative impact.

To lend further credibility to the assumption that selection of students across schools, due to the larger number of Asian students in the Dragon cohort, is not large enough to generate our results, we conduct three tests. In Panel E, we exclude from the sample areas in the top quintile of the 1990 Chinese population share, where moves to other neighborhoods may have been most likely for two reasons. First, due to the fixed costs of moving, non-Asian families in areas that experienced the largest shocks are disproportionately more likely to respond. Second, the increase in the share of Asian children and the negative impacts on tests scores may be more easily observable for parents of non-Asians students in highly-exposed areas, whereas they may be less obvious in areas with lower exposure. Reassuringly, the point estimates obtained from the sample excluding the most exposed areas are *larger* than in the baseline specification: the estimates are -2.175 for ELA ($p < 0.01$) and -0.842 for math ($p = 0.103$). In Appendix Figure A6, we also show that these results are robust across a wide range of 1990 Chinese population share cutoffs.

Another concern could be that parents may transfer their child to another school within the same residential area. Transfers between public schools are unlikely because only special needs and circumstances could enable students to move to an undesigned school.¹⁹ Children may, however, apply to a local charter or private school. Access to these schools is restricted by screening and, in case of private schools, high tuition fees. Nevertheless, in Panel F, we exclude from the sample public schools that are close to a charter or private school, because transfers can be expected to be more likely when an alternative school is located nearby.²⁰ When we drop the quintile of public schools that are closest to a charter or private school, the magnitude of the estimates is, again, larger than in the baseline specification and both estimates are significant at the 5% level. The estimates are little affected by the choice of the distance cutoff (Appendix Figure A7).

As a third test for student mobility, we examine student attrition by estimating the impact of the instrument on within-school changes in the number of non-Asian students from third to fourth and fourth to fifth grade.²¹ For instance, parents may try to move their child to another school when they observe a reduction in test scores. The results are provided in Appendix Table A7. For both outcomes, the estimates are small and statistically insignificant. Overall, the findings in Panels E and F and Appendix Table A7 suggest that selection of non-Asian students in response to the fertility shock in the Dragon year is unlikely to drive our results.

The fertility shock could affect the 2001 cohort, because the Chinese Dragon year ends on January 23, 2001. Furthermore, if spillovers across grades are important, children born in 1999 and 2001 may be most affected because of the smallest age difference relative to the Dragon cohort. In Panel G, we exclude these adjacent cohorts from the sample.

¹⁹ These criteria are: 1) medical reasons, 2) students' safety, 3) parent's employer being located far from the designated school, 4) a sibling attending a different school, and 5) own school being listed as a school in need of improvement or low-achieving school in the last two years.

²⁰ Addresses of charter and private schools are drawn from the DOE and Private Schools Universe Survey data. Following the same geocoding procedure as for public schools (see section 2.1), we assign coordinates to charter and private schools operating in NYC during the period of analysis.

²¹ We restrict this analysis to grades 3–5 because the majority of schools in our sample end in the fifth grade. Hence the number of observations for changes in enrollment across higher grades is too small for precise estimation.

The estimates for both math and ELA are similar compared to the baseline and suggest that the children born in the first 23 days of the year 2001 or spillovers to adjacent cohorts have little impact on our results.

Panel H shows results for a sample including grades 3–6, in which the Dragon cohort is observed. The point estimates are similar compared to our preferred baseline estimates obtained from the sample including all grades 3–8, although the precision of the estimation is lower. We prefer the specification including grades 3–8 as it provides smaller standard errors. This is because the larger sample contributes to the estimation of school fixed effects, which increases the power of the analysis.

Panel I shows results for a specification including the square of the 1990 Chinese population share interacted with the Dragon dummy as an additional instrument. This is motivated by the concave function of the theoretical impact in Eq. (1). Allowing for the nonlinear first-stage impact of the instrument does not affect the estimates appreciably.²²

4.3. Effects on performance levels

Columns 1–4 in Table 5 show estimates of the impacts of the share of Asian students on the share of non-Asian students in each of the four performance levels. Columns 5 and 6 show results for the share of non-Asian students at the two lowest and at the lowest and highest performance levels, respectively. The former specification estimates the effect on the share of students who are not meeting all learning standards while the latter tests for the dispersion of the performance level distribution. For all specifications, the first-stage estimate is highly significant (see Appendix Table A10).

For math, we detect an increase in the share of students at the lowest performance level (0.509 ; $p < 0.05$). The rise in the share of non-Asian students at the lowest performance level, combined with the significant reduction in their math scores (Table 3), suggests that the increase in the share of Asian students causes a larger fraction of non-Asian students to lag behind in this subject. The estimate means that a 10-percentage-point increase in the share of Asian students increases the share of non-Asian students who fail to meet the learning standards by around 5.1 percentage points or around 61% from the sample mean. The estimate in column 5 indicates a more general (marginal) reduction in non-Asian math proficiency.

The pattern of the estimates is slightly different for ELA, for which we observe a statistically significant reduction in the share of non-Asian students at the highest performance level

(-0.309 ; $p < 0.05$). We also detect a (marginally) significant increase for level 2, at which students only partially meet the learning standards (0.413 ; $p < 0.10$). The estimate in column 6 is consistent with the contraction of the non-Asian ELA performance distribution. Overall,

²² Appendix Table A8 shows results for alternative definitions of the instrument. These specifications are based on various Asian subpopulations (Chinese, Asian excluding Asian Indians, and all Asian) and school neighborhoods (census tracts within 500, 1,000, 2,000, and 3,000 meters of the school). The IV estimates for ELA range from -0.588 to -0.939 and are all statistically significant ($p < 0.05$). For math, the IV estimates range from -0.310 to -0.598 , with six of the 12 estimates being significant at the 5% or lower risk level and three being significant at the 10% risk level. The smallest buffer of 500 meters provides the smallest p-values. This is likely resulting from the fact that defining school neighborhoods by a wider radius leads to more overlap between school catchment areas. We cannot explicitly test for this hypothesis, however, as data on the actual school catchment areas are not available. Appendix Table A9 shows estimates for specifications using the share of Asian students in the school in the year 2006 (before the Dragon cohort enters the data) interacted with the Dragon dummy as the instrument. The IV estimates for this specification are also negative (-1.343 for ELA and -0.850 for math). We prefer the results based on the instruments constructed from the 1990 census data because these data allow us to focus on Asian subpopulations among which the Dragon belief is prevalent and are realized before the fertility shock in the year 2000.

Table 5
Effects of Asian Peers on the Share of Non-Asian Students by Performance Level.

Performance level:	(1) Level 1: Learning standards not met	(2) Level 2: Learning standards partially met	(3) Level 3: Learning standards met	(4) Level 4: Learning standards met with distinction	(5) Levels 1 and 2	(6) Levels 1 and 4
Math	0.509**	0.129	−0.329	−0.310	0.657*	0.173
Dependent Mean (%)	(0.250) 8.41	(0.225) 27.27	(0.205) 45.81	(0.291) 18.50	(0.367) 35.68	(0.258) 26.92
ELA	−0.268	0.413*	0.164	−0.309**	0.144	−0.578*
Dependent Mean (%)	(0.255) 9.47	(0.235) 38.56	(0.298) 47.48	(0.130) 4.49	(0.295) 48.02	(0.296) 13.96
Observations	49,872	49,872	49,872	49,872	49,872	49,872
Schools	1080	1080	1080	1080	1080	1080

Notes: This table reports IV estimates of the effect of the share of Asian students on the share of non-Asian students at the four performance levels separately for ELA and math, using the interaction term between the 1990 Chinese population share and the Dragon dummy as the instrument. Each cell reports a coefficient from a separate regression. The fifth and sixth columns show results for specifications using the share of students at the two lowest levels and of students at the lowest and highest levels as outcomes, respectively. All specifications control for grade-level enrollment and school, grade-by-year, and race fixed effects. Regressions are weighted by the number of students taking the test. Standard errors clustered at the census-tract level are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

compared to math, the impacts appear to be weaker at the lower tail of the ELA performance distribution. The estimates suggest that the decline in ELA test scores occurs across the three highest performance levels; the reduction in the top-performing group and the smaller increase in the second highest group suggest that some of these students move to lower performance levels. Moreover, the decline in the share of students at the lowest performance level suggests that some low-achieving students may benefit from the larger share of Asian students in this subject. Section 5 discusses some possible explanations for these findings.

4.4. Heterogeneity by race/ethnicity and grade

Panel A of Table 6 shows the impacts of the share of Asian students on non-Asian test scores by race. For math, the point estimates are negative for all groups, and (marginally) significant for the Hispanic group. The results for ELA are similar, except for the positive but insignificant point estimate for white students. For ELA, also the estimate for the group combining black and Hispanic students is (marginally) significant. Although caution is in order when interpreting these findings due to the relatively low precision of some estimates, and the lower precision of the first stage for the sample of black students (see Appendix Table A11), the results indicate fairly similar negative effects

Table 6
Effects of Asian Peers on Test Scores by Race/Ethnicity and Grade.

	(1) White	(2) Hispanic	(3) Black	(4) Hispanic and Black
Panel A: Race/ Ethnicity				
ELA	0.393	−0.509*	−1.213	−0.478*
	(0.456)	(0.294)	(0.837)	(0.268)
Math	−0.453	−0.518*	−0.773	−0.496
N	(0.401)	(0.308)	(0.998)	(0.325) 40,731
	9104	21,907	18,824	
Schools	501	1051	956	1077
Panel B: Grade				
ELA	−0.618*	−0.215	−0.883*	−0.500
	(0.320)	(0.470)	(0.501)	(0.532)
Math	−1.019**	−0.191	−0.466	−0.473
N	(0.402)	(0.394)	(0.403)	(0.723) 7216
	9982	9937	9930	
Schools	722	715	731	598

Notes: This table reports IV estimates of the effect of the share of Asian students on non-Asian ELA and math test scores by race/ethnicity and grade, using the interaction term between the 1990 Chinese population share and the Dragon dummy as the instrument. All specifications control for grade-level enrollment and school, grade-by-year, and race fixed effects. Regressions are weighted by the number of students taking the test. Standard errors clustered at the census-tract level are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

on math scores across non-Asian subgroups, whereas the negative impact on ELA scores appears to be driven by a reduction in achievement among Hispanic and black students.

Panel B reports IV estimates by grade. We report results for grades 3–6 as the sixth grade is the last one in which we observe the Dragon cohort. All first-stage coefficients are large and significant at the 5% risk level except for the sixth grade, for which the sample is the smallest (see Appendix Table A12). For both subjects, the IV estimates are negative for all grades and range from −1.019 to −0.191. The estimates for third grade scores are −1.019 ($p < 0.05$) for math and −0.618 ($p < 0.10$) for ELA. We also detect a (marginally) significant impact on fifth grade ELA score. Overall, although these estimates are based on smaller samples and the precision of the estimation reduces toward the sixth grade, they suggest that the negative effects persist across grades.

4.5. Heterogeneity by school characteristics

In this section, we examine whether the impact of Asian peers varies by school characteristics. We estimate the following reduced-form regression:

$$Y_{rsgrt} = \xi_1 CS_{s,1990} + \xi_2 Dragon_{grt} + \xi_3 CS_{s,1990} Dragon_{grt} + \xi_4 Z_s Dragon_{grt} + \xi_5 CS_{s,1990} Z_s + \xi_6 CS_{s,1990} Dragon_{grt} Z_s + \beta_1 X_{rsgrt} + u_{rsgrt} \quad (3)$$

Here Z_s is a school characteristic and the parameter of interest is the coefficient on its interaction with the instrument, ξ_6 . This coefficient tests whether the instrument has a different impact with respect to the school-level variable Z_s . The vector X_{rsgrt} includes enrollment at the grade and school, race, and grade-by-year fixed effects. As before, we weight the regressions by the number of students taking the test and cluster the standard errors at the census-tract level.

Table 7 reports the results. We start by examining whether the impacts are heterogeneous by the level of racial/ethnic fractionalization of the non-Asian student population. We use the ethno-linguistic fractionalization (ELF) index, $F_s = 1 - \sum_r r_{sr}^2$ (see, e.g., Bossert, D'Ambrosio, & Ferrara, 2011), where r_{sr} is the share of students in a non-Asian ethnic/racial subgroup r (white, Hispanic, or black) in school s in 2006.²³ Schools with a higher index of fractionalization have more similarly sized non-Asian subgroups. The Subculture model of peer effects

²³ This index measures the probability that two randomly selected non-Asian students belong to different racial/ethnic groups. When all non-Asian students in a school belong to the same racial/ethnic group, the index is equal to 0. For three non-Asian groups, the maximum value of the index is $1 - 3 \cdot (1/3)^2 \approx 0.77$.

Table 7
Heterogeneity.

	(1) ELA score	(2) Math score
A. 2006 Racial Fractionalization <i>N</i> = 47,679; Schools=966	−0.605** (0.242)	−0.133 (0.203)
B. 2006 Enrollment <i>N</i> = 47,679; Schools=966	0.00030* 0.00018	0.000070 0.000121
C. 2006 Fraction of Teachers with less than 3 Yrs. of Experience (%) <i>N</i> = 47,347; Schools=956	0.00123 (0.00325)	0.000893 (0.00391)
D. 2006 Pupils per Teacher Ratio <i>N</i> = 47,347; Schools=956	0.0113 (0.0166)	0.0136 (0.0132)
E. 2006 Number of Teachers <i>N</i> = 47,347; Schools=956	0.00320 (0.00231)	−0.00043 (0.00170)
F. 2007 Grade-Level Average Class Size <i>N</i> = 44,510; Schools=917	0.00318 (0.00838)	−0.0107 (0.00745)
G. Current Enrollment ≤ 25 <i>N</i> = 49,872; Schools=1080	0.0249 (0.105)	−0.434** (0.220)

Notes: This table reports reduced-form coefficients on the interaction between the instrument and a school characteristic indicated by the panel title. Each cell reports a coefficient from a separate regression. For example, the estimates in the first row are for the coefficient on the interaction between the 1990 Chinese population share, Dragon dummy, and index of racial fractionalization. All specifications control for the main effects and the interaction between the Dragon dummy and school characteristic, the interaction between the Dragon dummy and 1990 Chinese population share, and the interaction between the school characteristic and 1990 Chinese population share, and they include total enrollment and school, grade-by-year, and race fixed effects. The index of racial fractionalization, enrollment, pupils per teacher ratio, number of teachers, and percentage of teachers with less than three years of experience are measured in 2006. Class size is measured in 2007, which is the first year when it is available in our data. The dummy for enrollment less than or equal to 25 students is measured in the current year. Variation in the number of observations across specifications is due to the unavailability of data for some variables in some schools. For example, in Panel B, the estimation does not include schools for which enrollment in 2006 is not observed, such as schools that are observed for the first time after 2006. On the other hand, in Panel G, the interaction variable is based on current enrollment, which is available for all observations in the baseline sample. Regressions are weighted by the number of students taking the test. Standard errors clustered at the census-tract level are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

suggests that these schools may experience higher levels of cultural conflict (Hoxby & Weingarth, 2005). In the context of our study, pre-existing cultural conflicts between non-Asian subgroups could amplify the negative impacts of the change in the racial composition. The rise in the share of Asian students may also trigger cultural rejection, which could further increase conflict and reduce student achievement. Indeed, in Panel A, we detect negative coefficients on the interaction term between the instrument and the index of racial fractionalization. The estimate for ELA of -0.696 ($p < 0.01$) indicates that non-Asian ELA scores decrease more in schools with a more fractionalized non-Asian student population. This estimate means that increasing the fractionalization index by 0.10 points increases the impact of the instrument on non-Asian ELA scores by around 60% compared to the baseline reduced-form estimate in Table 3. For math, the estimate is also negative, but smaller and not statistically significant at the conventional confidence levels.

In Panels B-E, we show that the impact of Asian peers does not change appreciably with enrollment, teacher experience, pupils per teacher ratio, or number of teachers, all measured in 2006 before the Dragon cohort enters the school; we detect only one marginally significant coefficient, for enrollment. This estimate is positive and small and means that, if anything, the magnitude of the negative impact of Asian peers is larger in schools with lower pre-Dragon-cohort grade-level enrollment. These findings provide evidence that pre-existing differences in the quality and amount of teaching resources or larger grades are

unlikely to explain our results.²⁴

We then turn the focus to class size. Previous research has shown that larger classes can have negative effects on scholastic achievement (e.g., Angrist & Lavy, 1999; Chetty et al., 2011; Fredriksson, Öckert, & Oosterbeek, 2013). We have shown that our instrument does not affect class size and that controlling for it has little impact on our IV estimates. However, this does not rule out the possibility that the impacts of racial composition may vary with class size. We test for this hypothesis in Panel F, but find no evidence of a significant interaction effect.

In Panel G, we explore whether tracking of students to classes can explain our results. To do so, we examine possible differential effects in grades where enrollment is below or equal to the upper limit of 25 for class size: In such grades, there is likely to be only one class and hence no tracking. The coefficient on the interaction term is negative and significant for math ($p < 0.05$); for ELA we do not reject the hypothesis of similar effects. These findings suggest that tracking of students to classes is unlikely to explain our results; if tracking was driving our results, we should find larger negative effects for larger grades where tracking is possible, but the results do not indicate this. A potential explanation for the larger effect on math in grades with 25 or fewer students could be that in smaller grades the interaction between classmates is more intensive, and interaction with peers in other classes in the same grade is limited or not possible.

5. Discussion

The parameter recovered by our IV strategy is the Local Average Treatment Effect (LATE). It identifies the impact of the increase in the share of Asian students due to the additional Asian students who are born in the Dragon year as a result of the belief that Dragon children are luckier, brighter, and more likely to succeed in life. Previous research using U.S. data has shown that Asian mothers of children born in a Dragon year are, on average, more educated and have higher income than other Asian mothers (Johnson & Nye, 2011). In their research using Chinese data, Mocan and Yu (2017) show that Chinese parents of Dragon children have higher expectations for their children's education and invest more time and financial resources in them compared to other parents, but the Chinese Dragon children do not have higher self-esteem or expectations about the future. In line with their findings, we also find that Asian Dragon students perform better in school compared to Asian non-Dragon students. However, our instrumental variable does not induce a change in the average Asian test scores, although we find some evidence of an increase in the dispersion of the performance level distribution. This means that an average additional Asian student in the Dragon cohort is similar in terms of scholastic performance compared to an average Asian student who would be born in the Dragon year in the absence of the Dragon belief. We conclude that while changes in the average Asian peer achievement do not contribute to our IV estimates, the Asian peer group, which is the source of the estimated peer effects in our study, has 0.046σ and 0.094σ higher average test scores in ELA and math, respectively, compared to other Asian American student in NYC public schools.²⁵

We next compare our findings to other studies that have examined the effects of racial group composition on student achievement. Angrist and Lang (2004) find little evidence of socially or statistically significant overall effects of Metco students on their non-Metco classmates.

²⁴ We provide further evidence supporting the interpretation that teaching resources are unlikely to affect our results by estimating the impact of the Dragon cohort entering a school on the number of teachers, pupils per teacher ratio, and the fraction of teachers with less than three years of experience in a school-level regression (for details, see appendix A.5). We find no statistically significant effects on these outcomes (see Appendix Table A13).

²⁵ These numbers are calculated by dividing the estimates in columns 1 and 3 of Appendix Table A2 with the New York State test score standard errors.

However, for black non-Metco third graders, they find that a 10-percentage-point increase in the share of Metco students reduces their reading and language test scores by around 0.6σ . The estimates in Hanushek et al. (2009) imply that a 10-percentage-point increase in the proportion of black students reduces black achievement by around 0.02σ and white achievement by around 0.01σ . Hoxby (2000) finds that for a 10-percentage-point increase in the share of black students, reading scores reduce by 0.06 to 0.25σ , and math scores reduce by 0.04 to 0.19σ . The magnitude of our estimates, suggesting that a 10-percentage-point increase in the share of Asian students decreases non-Asian test scores by around 0.16σ in ELA and 0.14σ in math, falls within the range of estimates identified in these studies.

As discussed above, Asian peers in our analysis are on average top-achievers. Therefore, it is also useful to compare our results to previous findings on the impacts of high-achieving peers on educational outcomes. Because peer effects can vary by the level of education (Sacerdote, 2011), we focus on results from studies that have examined students in primary and secondary education. Hanushek et al. (2003) and Vigdor and Nechyba (2007) find significant positive impacts of peer achievement on own achievement in public schools in Texas and North Carolina. Lavy, Silva, and Weinhardt (2012) estimate within-pupil regressions for secondary school students in the UK. They find little evidence of overall impacts of high-achieving peers, but their results suggest that girls benefit from interactions with very bright peers. Gibbons and Telhaj (2016) study re-sorting of students when they move from primary to secondary schools in the UK. They find small positive impacts of peer achievement on test scores. Unlike these previous studies, we find negative average effects of high-achieving Asian students on their schoolmates' test scores. This points to the possibility that, in our context, racial dynamics may be a stronger determinant of educational outcomes than peer achievement.²⁶

Like many other studies in the field, we are unable to explicitly test for the specific mechanisms of racial group effects due to the unavailability of data on teaching practices and student behavior. Nevertheless, we provide evidence against several channels through which the increase in the share of Asian students could affect non-Asian test scores (congestion, changes in class size, tracking, attrition, moves to and from private and charter schools, and school responses). We believe that the most plausible explanations for the negative peer effects that we detect are student discouragement and teacher responses.

For ELA, the contraction of the performance distribution could be the result of teacher responses stemming from the fact that a large fraction of the Asian population is bilingual.²⁷ Bilingual Asian children can experience delays in acquiring some formal aspects of English, such as vocabulary, even if their parents are fluent in English (Bialystok, Luk, Peets, & Yang, 2010). Moreover, Asian parents have typically high expectations for their children's education (Chao & Tseng, 2002; Mocan & Yu, 2017). For these reasons, Asian students could require more teacher's attention to achieve their study goals. This may reduce the pace of ELA instruction, which could benefit low-achieving students but slow

down the progress of high-achieving students, leading to the contraction of the ELA performance distribution (Table 5). Moreover, this can occur simultaneously with the reduction in ELA test scores (Table 3) if the increase in achievement among low-achieving students does not offset the decline in achievement among high-achieving students.

For math, the negative effects on the share of non-Asian students in the two highest performance levels could result from discouragement because Asian students are performing extremely well in this subject. Discouragement effects may be reinforced by stereotypical perceptions of teachers that vary across minority groups. Previous research has shown that Asian students are often viewed more positively by teachers than white students (McGrady & Reynolds, 2013) and Asian American children are sometimes held to a "model student" stereotype (Rose-nbloom & Way, 2004; Wong, 1980).²⁸ We also note that the rise in the share of non-Asian students at the lowest math performance level could be caused by teachers increasing the pace and coverage of instruction to better suit Asian students. This could be particularly harmful for students who are already experiencing difficulties in this subject.

We believe that the observed differences in the effects across the performance levels between math and ELA are not unexpected because Asian students are particularly well achieving in math and achievement in ELA is more dependent on the children's language environment.

6. Concluding remarks

This study examines the impact of Asian students on the scholastic performance of their schoolmates. We employ test score data for public schools in NYC, which houses one of the largest Asian populations in the U.S. We address endogeneity concerns by exploiting plausibly exogenous variation in the share of Asian students within schools and across cohorts as a result of the fertility spike among the Asian populations in the Chinese year of the Dragon. Our identification strategy exploits the fact that the relative magnitude of the fertility shock varies geographically with the local historical share of the Chinese population.

Our study contributes to the literature on the educational impacts of racial group composition, and more generally, on peer effects in education. We provide new evidence on the effects of Asian peers on test scores in a setting that combines data from a school district with a large Asian population and considerable quasi-experimental variation in the share of Asian students. We find that exposure to Asian peers has a negative causal effect on both the ELA and math scores of non-Asian students. These negative impacts lead to an increase in the share of non-Asian students who fail to demonstrate the skills expected at the grade, especially in math.

Past studies on the effects of racial composition in schools have nearly always focused on the share of black students and have often found negative impacts on educational outcomes. The strong focus on black peers is motivated by their relatively low average achievement, existing black-white educational achievement gaps, and historically high levels of black-white school segregation. Yet, the strong focus on one student group might mask other important racial dynamics within schools. By moving beyond black-white segregation, our study shows that an increase in the share of a well-achieving minority group can have negative impacts on average student achievement, suggesting that racial composition (vis-à-vis peer achievement) has an independent and important role in determining educational outcomes. This interpretation

²⁶ Several studies have estimated peer effects in colleges and universities exploiting a random or quasi-random assignment of peer groups. Sacerdote (2001) examines the effects of randomly assigned roommates finding that peers have a modest impact on academic performance. Zimmerman (2003) finds that students are negatively affected by being assigned a roommate in the lowest 15% of the achievement distribution. Carrell, Fullerton, and West (2009) find positive effects of peer achievement on math and science grades. Feld and Zölitz (2017) find that, on average, students benefit by a small amount from being exposed to better peers, but the test scores of low-achieving students decline when they are exposed to high-achieving peers. Golsteyn, Non, and Zölitz (2018) provide evidence of positive impacts of persistent peers on academic achievement.

²⁷ In 2011-2015, the fraction of Asians who are bilingual is 60% among individuals aged 5-18 and 45% among individuals aged 16-64 in the U.S. (Chiswick and Gindelsky, 2016; Ee, 2019).

²⁸ For other evidence of the significance of racial dynamics between teachers and students, see also Dee (2004), who finds evidence of significant racial teacher-student mismatch effects on test scores for white and black students in a randomized experiment, and Dee (2005), who finds that teachers' perceptions are more positive for students who share the same racial designation. In their study on special education identification in Florida schools, Elder et al. (forthcoming) find that black students are over-identified in schools with relatively low share of minorities.

is consistent with a body of literature suggesting that interactions between teachers and students (or their parents) are affected by racial/ethnic backgrounds, and that racial mismatch may complicate classroom interactions and undermine academic achievement (e.g., Dee, 2004, 2005; Lareau & Weininger, 2003; Valdes 1996).

From a methodological perspective, we contribute to the literature by offering a new approach to the identification of group composition effects based on explicit fertility shocks due to cultural beliefs. Our identification strategy is not specific to the schooling context. Therefore, it can be helpful for future studies examining the impacts of racial composition on other educational, economic, and social outcomes. Our findings have some important implications for school management as well. They show that a change in the racial composition due to a local population shock can have substantial, and sometimes unexpected, consequences for student achievement. Our finding of the significant increase in the share of students who lag behind further emphasizes the importance of school policies that attenuate the potentially adverse impacts of such shocks. Establishing the mechanisms through which racial composition affects student achievement to help determine such policies is a task for future research.

Data availability statement

The data used in this article are available online. Links to data sources are provided in Appendix A.

Disclosure statement

The authors declare that they have no relevant or material financial interests that relate to the research described in this paper and that they have not received relevant supporting funds.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.econedurev.2021.102129.

References

- Andrews, Isaiah, Stock, James H., & Sun, Liyang (2019). Weak Instruments in Instrumental Variables Regression: Theory and Practice. *Annual Review of Economics*, 11, 727–753.
- Angrist, Joshua D. (2014). The perils of peer effects. *Labour Economics*, 30, 98–108.
- Angrist, Joshua D., & Lang, Kevin (2004). Does school integration generate peer effects? Evidence from Boston's Metco Program. *American Economic Review*, 94(5), 1613–1634.
- Angrist, Joshua D., & Lavy, Victor (1999). Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *Quarterly Journal of Economics*, 114(2), 533–575.
- Angrist, Joshua D., & Pischke, Jörn-Steffen (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton university press.
- Bialystok, Ellen, Luk, Gigi, Peets, Kathleen F., & Yang, Sujin (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*, 13(4), 525–531.
- Bossert, Walter, D'Ambrosio, Conchita, & Ferrara, Eliana La (2011). A Generalized Index of Fractionalization. *Economica*, 78(312), 723–750.
- Budiman, Abby, Cilluffo, Anthony, & Ruiz, Neil G. (2019). "Key facts about Asian origin groups in the U.S." Pew Research Centre.
- Carrell, Scott E., Fullerton, Richard L., & West, James E. (2009). Does your cohort matter? Measuring peer effects in college achievement. *Journal of Labor Economics*, 27(3), 439–464.
- Carrell, Scott E., Hoekstra, Mark, & Kuka, Elira (2018). The long-run effects of disruptive peers. *American Economic Review*, 108(11), 3377–3415.
- Chao, Ruth, & Tseng, Vivian (2002). Parenting of Asians" in Handbook of Parenting. In *Social conditions and applied parenting*, 4. edited by Marc H. Bornstein. Taylor and Francis.
- Chetty, Raj, Friedman, John N., Hilger, Nathaniel, Saez, Emmanuel, Schanzenbach, Diane Whitmore, & Yagan, Danny (2011). How does your kindergarten classroom affect your earnings? Evidence from Project STAR. *Quarterly Journal of Economics*, 126(4), 1593–1660.
- Chiswick, Barry R., & Gindelsky, Marina (2016). Determinants of bilingualism among children: An econometric analysis. *Review of Economics of the Household*, 14(3), 489–506.
- Dee, Thomas S. (2004). Teachers, race, and student achievement in a randomized experiment. *Review of Economics and Statistics*, 86(1), 195–210.
- Dee, Thomas S. (2005). A teacher like me: Does race, ethnicity, or gender matter? *American Economic Review*, 95(2), 158–165.
- Duflo, Esther, Dupas, Pascaline, & Kremer, Michael (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *American Economic Review*, 101(5), 1739–1774.
- Ee, Jongyeon. (2019). Bamboo bridges or barriers? Exploring advantages of bilingualism among Asians in the US labor market through the lens of superdiversity. *Bilingual Research Journal*, 42(2), 252–268.
- Feld, Jan, & Zölitz, Ulf (2017). Understanding peer effects: On the nature, estimation, and channels of peer effects. *Journal of Labor Economics*, 35(2), 387–428.
- Fredriksson, Peter, Öckert, Björn, & Oosterbeek, Hessel (2013). Long-term effects of class size. *Quarterly Journal of Economics*, 128(1), 249–285.
- Gibbons, Stephen, & Telhaj, Shqiponja (2016). Peer Effects: Evidence from Secondary School Transition in England. *Oxford Bulletin of Economics and Statistics*, 78(4), 548–575.
- Golsteyn, B. H., Non, A., & Zölitz, U. (2018). The impact of peer personality on academic achievement. *Journal of Political Economy*, 129(4), 1052–1099.
- Goodkind, Daniel M. (1995). The significance of demographic triviality: Minority status and zodiacal fertility timing among Chinese Malaysians. *Population Studies*, 49(1), 45–55.
- Hanushek, Eric A., Kain, John F., Markman, Jacob M., & Rivkin, Steven G. (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics*, 18(5), 527–544.
- Hanushek, Eric A., Kain, John F., & Rivkin, Steven G. (2009). New evidence about *Brown v. Board of Education*: The complex effects of school racial composition on achievement. *Journal of Labor Economics*, 27(3), 349–383.
- Hoxby, Caroline M. (2000). "Peer effects in the classroom: Learning from gender and race variation." NBER Working Paper no. 7867.
- Hoxby, Caroline M., & Weingarth, Gretchen (2005). Taking race out of the equation: School reassignment and the structure of peer effects. *Unpublished manuscript*.
- Imberman, Scott A., Kugler, Adriana D., & Sacerdote, Bruce I. (2012). Katrina's children: Evidence on the structure of peer effects from hurricane evacuees. *American Economic Review*, 102(5), 2048–2082.
- Johnson, Noel D., & Nye, John V. C. (2011). Does fortune favor dragons? *Journal of Economic Behavior & Organization*, 78(1), 85–97.
- Lareau, A., & Weininger, E. B. (2003). Cultural capital in educational research: A critical assessment. *Theory and Society*, 32(5–6), 567–606.
- Lavy, Victor, Silva, Olmo, & Weinhardt, Felix (2012). The good, the bad, and the average: Evidence on ability peer effects in schools. *Journal of Labor Economics*, 30(2), 367–414.
- Lazear, Edward P. (2001). Educational production. *Quarterly Journal of Economics*, 116(3), 777–803.
- Manski, Charles F. (1993). Identification of endogenous social effects: The reflection problem. *Review of Economic Studies*, 60(3), 531–542.
- McGrady, Patrick B., & Reynolds, John R. (2013). Racial mismatch in the classroom: Beyond black-white differences. *Sociology of Education*, 86(1), 3–17.
- Mocan, Naci H., & Yu, Han (2017). "Can superstition create a self-fulfilling prophecy? School outcomes of Dragon children of China." NBER Working Paper no. 23709.
- Rivkin, Steven, & Welch, Finis (2006). Has school desegregation improved academic and economic outcomes for blacks? In *Handbooks of Economic of Education*. Amsterdam: North Holland, 2 (pp. 1019–1049). edited by Eric A. Hanushek and Finis Welch.
- Rogers, Todd, & Feller, Avi (2016). Discouraged by peer excellence: Exposure to exemplary peer performance causes quitting. *Psychological Science*, 27(3), 365–374.
- Rosenbloom, Susan Rakosi, & Way, Niobe (2004). Experiences of discrimination among African American, Asian American, and Latino adolescents in an urban high school. *Youth & Society*, 35(4), 420–451.
- Sacerdote, Bruce. (2001). Peer effects with random assignment: Results for Dartmouth roommates. *Quarterly Journal of Economics*, 116(2), 681–704.
- Sacerdote, Bruce. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far? *Handbook of the Economics of Education*, 3(3), 249–277.
- Skeels, Christopher L., & Windmeijer, Frank (2018). On the Stock-Yogo Tables. *Econometrics*, 6(4), 44.
- Vigdor, Jacob, & Nechyba, Thomas (2007). Peer effects in North Carolina public schools. In L. Woessmann, & P. E. Peterson (Eds.), *Schools and the equal opportunity problem* (pp. 73–101). Cambridge, MA: The MIT Press.
- Welch, Finis, & Light, Audrey (1987). *New evidence on school desegregation*. Washington, DC: U.S. Commission on Civil Rights.
- Wong, Morrison G. (1980). Model students? Teachers' perceptions and expectations of their Asian and White students. *Sociology of Education*, 53(4), 236–246.
- Yip, Paul S. F., Lee, Joseph, & Cheung, Yin-Bun (2002). The influence of the Chinese zodiac on fertility in Hong Kong SAR. *Social Science & Medicine*, 55(10), 1803–1812.
- Zimmerman, David J. (2003). Peer effects in academic outcomes: Evidence from a natural experiment. *Review of Economics and Statistics*, 85(1), 9–23.